Contribution ID: **333**                                                         Type: **Talk**

# Reading Tea Leaves - Understanding internal events and addressing performance issues within a CephFS/XRootD Storage Element.

*Thursday 24 October 2024 17:09 (18 minutes)*

Erasure-coded storage systems based on Ceph have become a mainstay within UK Grid sites as a means of providing bulk data storage whilst maintaining a good balance between data safety and space efficiency. A favoured deployment, as used at the Lancaster Tier-2 WLCG site, is to use CephFS mounted on frontend XRootD gateways as a means of presenting this storage to grid users.

These storage systems are complex and self-correcting, but despite access to a myriad of metrics the inner workings of the storage tend to be opaque to the storage admin. One of the common problems seen within Ceph based systems are "Slow Ops"- instances of operations that take longer than expected, that are also often blocking in nature, impacting the overall performance and reliability of the system. These could be caused by, for example, intensive client side usage, internal CEPH data movement or hardware and/or network issues. Identifying the causes of a slow operation can provide a means to prevent or reduce the impact of future occurrences, leading to an increase in performance and reliability.

We detail the Lancaster Grid Site's attempts to understand the causes of and mitigate against these "Slow Ops"and other performance bottlenecks within our storage system, with a focus on deletions as a case study on operations with a potential high-impact for the Ceph backend. We endeavour to bring together a holistic monitoring model, utilising Ceph metrics, detailed XRootD monitoring streams and client-side logging, in order to understand how data-management events impact the health of the storage.

**Authors:** HAND, Gerard (Lancaster University (GB)); DOIDGE, Matt (Lancaster University (GB)); LOVE, Peter (Lancaster University (GB)); SIMPSON, Steven (Lancaster University)

**Presenter:** DOIDGE, Matt (Lancaster University (GB))

**Session Classification:** Parallel (Track 1)

**Track Classification:** Track 1 - Data and Metadata Organization, Management and Access